<center>REMARKS</center>

The comments of the applicant below are each preceded by related comments of the examiner (in small, bold type).

The applicant thanks the examiner for the interview with the applicant's representatives on October 16, 2008. During the interview the patentability of proposed amended claims, similar to the amended claims presented here, was discussed.

> **The information disclosure statement filed 2/14/08 fails to comply with 37 CFR 1.98(a)(I), which requires the following: (1) a list of all patents, publications, applications, or other information submitted for consideration by the Office; (2) U.S. patents and U.S. patent application publications listed in a section separately from citations of other documents; and (3) the application number of the application in which the information disclosure statement is being submitted on each page of the list. The NPL listed in the information disclosure statement has been placed in the application file, but the information referred to by the NPL listed has not been considered. The NPL listed in the information disclosure statement is not a printed publication readily available to the public. Furthermore, the provided numbers correspond to applications which are still pending.**

A new form 1449 indicating the publication dates of the patent applications is being submitted. The claims and PAIR reports are "other information submitted for consideration by the Office" and are publicly available on the patent office website, although the patent owner is not sure of the dates on which they were first available publicly.

> **Claims recite "concordance scores". The specification defines "concordance scores" as "area under curve": "Clicking on the comparative model statistics button compares the concordance (area under ROC Curve) for both the sample dataset and the validation dataset as shown in FIGS. 258 and 25C." (see page 37, lines 5-7 and Figs. 258 and 25C). "The concordance statistic (c) [area under curve] of the validation dataset ..." (see page 37, line 9). The claims reciting "concordance scores" were interpreted as "area under curve".**

The term "concordance scores" can also be interpreted other than to be "area under curve".

> **Claims 1-1 9, 22-30, and 34-37, are rejected under 35 U.S.C. 103(a) as being unpatentable over by Cabena et al., (Cabena hereinafter), Intelligent Miner for Data Applications Guide (see IDS dated 12/18/06), taken in view of Harrison, (Harrison hereinafter), An Intelligent Business Forecasting System (see IDS dated 12/18/06).**

As to claim 1, Cabena discloses a machine-based method comprising in connection with a project in which a user generates a predictive model based on historical data about a system being modeled (see chapter 1.5.1, Pages 9-1 1): selecting variables having at least a first predetermined level of significance from a pool of potential predictor variables associated with the data, to form a population of predictor variables (see page 101, 2nd and 3rd paragraphs), extending the population of predictor variables to include non-linear interactions of variables (see page 93, 2nd paragraph) and extending the population of predictor variables to include linear and non-linear extensions with remaining previously excluded variables (see "supplementary variables" in "All other discrete and categorical variables and some interesting continuous variables were input as supplementary variables to be profiled with the clusters but not used to define them. These supplementary variables can be used to interpret the cluster as well. The ability to add supplementary variables at the outset of clustering is a very useful feature of Intelligent Miner, which allows the direct interpretation of clusters using other data very quickly and easily" in page 48, 1st paragraph), generating a possible model of the extended population of predictor variables using a subsample of the data by the model generation method (see "Feature Selection" and "Train and Test" in page 95), determining whether the possible model generalizes to the data other than the subsample (see page 101, last paragraph), applying the possible model to all of the data to generate a final model, cross-validating the final model using random portions of the data (see page 97, last paragraph), and interacting with the system being modeled based on the final model (see "To ensure that the model has not overfit the data and to assess the model performance against a data set that has the same characteristics as the application universe, the model should be executed against the test data in test mode" in page 102, 1st paragraph, lines 1-5 and "After having iteratively improved the models, you chose the best model" in page 102, 3rd paragraph, line 1).

While Cabena discloses generating a predictive model based on historical data about a system being modeled, Cabena fails to disclose automatically selecting a model generation method from among a set of available model generation methods to match characteristics of the historical data.

Harrison discloses automatically selecting a model generation method from among a set of available model generation methods to match characteristics of the historical data (see page 233, col. 2, next to last paragraph, last 7 lines).

Cabena and Harrison are analogous art because they are both related to predictive modelling.

Therefore, it would have been obvious to one of ordinary skill in this art at the time of invention by applicant to utilize the automatic model selection of Harrison in the method of Cabena because Harrison explore the possibility of the integration of expert systems technology with a forecasting decision support system (see page 229, col. 1, lines 1+), and as a result, Harrison reports that testing of his prototype shows that the system is useful for managers who have no forecasting technique and computing background and want to improve their decision making by means of quantitative forecasting (see page 235, col. 2, next to last paragraph).

Cabena's supplemental variables are discrete and categorical variables that are not used for defining his clusters and are not "cross products of at least two variables each within the first population of predictor variables" (specification, e.g., page 16, lines 27-29 and page 17, lines 24-27), as recited by amended claim 1. Nor are Cabena's supplemental variables "cross products of at least two variables, at least one of the variables being from the first population of predictor variables and having less than the first predetermined level of significance" (specification, e.g., page 16, lines 27-29 and page 18, lines 4-11), also recited by amended claim 1. Accordingly, Cabena did not describe and would not have made obvious the method recited by amended claim 1. Harrison has nothing to do with expanding the population of variables, let alone expanding by including cross products of variables.

> As to claim 34, Cabena discloses a machine-based method comprising receiving from separate sources, sets of potential predictor and dependent variables representing historical data about a system being modeled (see page 92, paragraphs 2-5), and enabling a user of a model generation tool to combine based on the dependent variables from the sets of potential predictor and dependent variables (see "Okay Customer Set", "Good Customer Set" and "Create Objective Variable" items in page 90, Fig. 46) to generate a model to be used in interacting with the system being modeled (see "To ensure that the model has not overfit the data and to assess the model performance against a data set that has the same characteristics as the application universe, the model should be executed against the test data in test mode" in page 102, 1st paragraph, lines 1-5 and "After having iteratively improved the models, you chose the best model" in page 102, 3rd paragraph, line 1). While Cabena discloses combining based on the dependent variables from the sets of potential predictor and dependent variables, Cabena fails to disclose combining at least two models. Harrison discloses combining at least two models (see page 233, col. 2, next to last paragraph, lines 11-15 and last paragraph).

As the examiner conceded, Cabena did not describe and would not have made obvious combining at least two models, as recited by claim 34. Harrison did not combine two models based on "response propensities of each model" in order to create cross-modal deciles, let alone based on "data weaving" to provide cross-modal optimization or concatenating the predictions of the two models in combing (specification, e.g., page 22, lines 6-10 and 12-16), as recited by amended claim 34.

When describing combining models, Harrison simply said "after the expert system suggests the appropriate models, ... let the system perform the models automatically, including

combining two models" (page 233, left column), but did not describe and would not have made obvious how to "combine two models". "Response propensities" as recited in claim 34 can be, for example, predictive outputs from each model such as probabilities of discrete responses (e.g., yes or no) or continuous variations (e.g., revenue being higher than an amount R or probability higher than a percentage p). "Data weaving" can be, for example, filtering of historical data for each model based on collective standards (e.g., level of significance of corresponding variables selected for each model). After reading Harrison, one skilled in the art would not have known Harrison's models are combined based on "response propensities" and "data weaving". One skilled would not have known how to combine Cabena's models as Harrison suggested, because one would not have known how to do so. Even if one skilled would have tried to combine Cabena's models, one would not have known to combine the models based on "response propensities" and "data weaving", let alone "concatenating the predictions of the models", also recited by claim 34.

Claims 2-11, 13-19, 22-23, 25-30, and 35-37 are patentable for at least the reasons discussed with respect to claims 1 and 34, from which they depend.

> Claims 31-33 and 38-40, are rejected under 35 U.S.C. 103(a) as being unpatentable over by Cabena taken in view of Harrison.
> As to claim 31, Cabena discloses a machine-based method comprising in connection with a project, generating a predictive model based on the historical data (see chapter 1.5.1, Pages 9-1 I), and displaying to a user a lift chart (see page 101, last paragraph, lines 1-5 and page 105, 1st and 2nd paragraphs), monotonicity (see page 101, last paragraph, last 3 lines and page 119, 2nd bullet from the bottom), and concordance scores (see Chapter 1.5.1, Pages 9-1 I) associated with each step in a step-wise model fitting process (see page 98, 2nd paragraph). While Cabena discloses generating a predictive model based on historical data about a system being modeled, Cabena fails to disclose automatically selecting a model generation method from among a set of available model generation methods to match characteristics of the historical data. Harrison discloses automatically selecting a model generation method from among a set of available model generation methods to match characteristics of the historical data about a system being modeled (see page 233, col. 2, next to last paragraph, last 7 lines).
> Cabena does not disclose expressly displaying to a user concordance scores. Examiner notes that the claims reciting "concordance scores" were interpreted as "area under curve".
> Official notice is taken that, displaying to a user the area under a curve was well known at the time the invention was made in the analogous art of Chapman et al., (Chapman hereinafter), CRISP-DM 1.0 Step-by-step Data Mining Guide, (see IDS dated 12/18/06). (See "area under a curve" in

"Outputs. Derived Attributes. Derived attributes are new attributes that are constructed from one or more existing attributes in the same record. An example might be area=length * width" in page 50, 2nd paragraph and "displaying to a user " in "The outputs produced during the Modeling phase can be combined into one report" in page 66, 1st paragraph, line 1).As a matter of fact, Examiner notes that "displaying to a user the area under a curve" is trivial to any engineer.

At the time the invention was made, it would have been obvious to a person of ordinary skill in the art to display to a user the area under a curve.

The suggestion/motivation to do so would have been to provide an additional visualization option (see "Results Visualization. The output of the clustering algorithms is an output data set and a visualization. The visual results display the number of clusters, the size of each cluster, the distribution of each variable in each cluster, and the importance of each variable to the definition of each cluster" in Cabena page 48, 3rd paragraph).

Therefore, it would have been obvious to modify Cabena to obtain the invention as specified in claims.

As the examiner implied, neither Cabena nor Harrison described or would have made obvious displaying to a user "concordance scores", as recited by claim 31. The examiner took official notice that the display of "concordance scores" to a user is obvious based on Chapman. However, "area" in Chapman is a derived attribute from existing attributes "length" and "width" for a record (page 49, "Task" and page 50, "Output"), and has nothing to do with "concordance scores" obtained based on a "receiver-operator-characteristic curve" and indicating to the user "goodness of fit" of the historical data to the generated predictive model (specification, e.g., page 10, lines 6-10 and page 35, lines 24-25), as recited by amended claim 31.

It would not have been obvious for one skilled in the art to display to a user in "concordance scores". If the examiner is to maintain this rejection, the applicant requests that new evidence be shown in support of the examiner's official notice.

Claims 32-33 are patentable for at least the reasons discussed with respect to claim 31. Claims 38-40 are amended to be dependent from claim 34 and are patentable for at least the reasons discussed with respect to claim 34.

All of the dependent claims are patentable for at least the reasons for which the claims on which they depend are patentable.

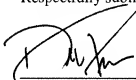Canceled claims, if any, have been canceled without prejudice or disclaimer.

Any circumstance in which the applicant has (a) addressed certain comments of the examiner does not mean that the applicant concedes other comments of the examiner, (b) made arguments for the patentability of some claims does not mean that there are not other good reasons for patentability of those claims and other claims, or (c) amended or canceled a claim does not mean that the applicant concedes any of the examiner's positions with respect to that claim or other claims.

Please apply $555 for the Petition for Extension of Time fee and any other charges or credits to deposit account no. 06-1050, referencing attorney docket no. 17146-0007001.

Respectfully submitted,

Date: _____10/28/8_____

David L. Feigenbaum
Reg. No. 30,378

Fish & Richardson P.C.
225 Franklin Street
Boston, MA 02110
Telephone: (617) 542-5070
Facsimile: (617) 542-8906

22048442.doc